A HVS based Perceptual Quality Estimation Measure for Color Images

Prateek Gupta[1], Priyanka Srivastava[1], Satyam Bhardwaj[1] and Vikrant Bhateja[1]
[1]Department of Electronics and Communication Engineering
Shri Ramswaroop Memorial Group of Professional Colleges
Lucknow-227105 (U.P), India
Email: {enggprateek.gupta, priyanka.srivastava4989,
er.satyam123, bhateja.vikrant}@gmail.com

***Abstract*—Human eyes are the best evaluation model for assessing the image quality as they are the ultimate receivers in numerous image processing applications. Mean squared error (MSE) and peak signal-to-noise ratio (PSNR) are the two most common full-reference measures for objective assessment of the image quality. These are well known for their computational simplicity and applicability for optimization purposes, but somehow fail to correlate with the Human Visual System (HVS) characteristics. In this paper a novel HVS based perceptual quality estimation measure for color images is proposed. The effect of error, structural distortion and edge distortion have been taken in account in order to determine the perceptual quality of the image contaminated with various types of distortions like noises, blurring, compression, contrast stretching and rotation. Subjective evaluation using Difference Mean Opinion Score (*DMOS*), is also performed for assessment of the perceived image quality. As depicted by the correlation values, the proposed quality estimation measure proves to be an efficient HVS based quality index. The comparisons in results also show better performance than conventional PSNR and Structural Similarity (SSIM).**

***Index Terms*—PSNR, HVS, RGB color model, Quality Measure (*Q*), Subjective Evaluation, *DMOS*.**

I. Introduction

Although the field of digital image processing is built on the foundation of mathematical formulations, yet human intuition and analysis plays a central role in the choice of one technique over another [1]. The visual quality of a color image is adjudged in the best way according to the look and feel of the image by the human visual system and hence it has become a major platform for the quality evaluation of color images. Distortions are introduced in an image during the process of storage, reconstruction, compression and enhancement. Therefore, assessment of quality is a crucial task in image processing applications [2]. Objective evaluation models use the mathematical expressions for assessment of image quality while the subjective models are based on physiological and psychological perception of human. Subjective models, evaluate the quality of the image as perceived by an individual observer. The features of both these models are incorporated in the HVS model that correlates well with the perceived image quality. The scope of conventional HVS based evaluation model is limited, as they consider error sensitivity as the only parameter for assessment of image quality [3]. The conventional indexes of quality assessment MSE and PSNR are operationally simple and have clear physical meaning but are unable to assess the similarity between different distortion types [4]. Compatibility of SSIM [5] with HVS characteristics accounts for its wide applicability, yet the performance of this metric degrade for poor quality and high texture images. Apart from its computational complexity (in comparison to MSE), SSIM is also not efficient for relative translations, scaling and rotation of images [4], [6]. Eric Wharton *et al.* [7] included Michelson contrast and MSE for making their proposed metrics more compatible to HVS. But their metrics does not account for structural and edge changes which are mainly noticed by human eyes. Wan Yang *et al.* [8] used a method of image quality assessment based on Region of Interest (ROI) and structural similarity. They had modified classical SSIM by using circular symmetric Gaussian weighting function and used gradient operator to enhance the performance of their metric. But their metric does not account for the edge modifications in the image which are mainly noticed by human eyes. Wei Fu *et al.* proposed a similarity index [9] for color images and used edge, luminance and structural similarity for the assessment of quality. However their model neglects the effects of other color image components like contrast that give information about the visual quality of image. Bo Wang *et al.* in their quality index [10] employ the HVS characteristics both in frequency and spatial domain. Ho-Sung Han *et al.* [11] used the gradient information for the assessment of image quality. The concept used in their work has been considering only the effect of large differences between the pixels of the original and distorted images. However, distortions projecting small differences in the pixels are neglected. Recently, Chen Yutuo *et al.* proposed an evaluation method for coding color images [12] based on the HVS model. Image pixels, region construction and edges are used as evaluation parameters in their work. The simulations performed by the authors have considered only the limited set of distortions, being silent about the effect of different noise types, rotation, blurring etc. Many objective evaluations proposed in the literature are at times proved complex and could not compete over the conventional PSNR [13]-[15]. Although, it is true that images with higher values of PSNR do not often yield good visualization by human

DOI: 01.IJSIP.03.01.11

observers. They are limited in approach to assess over enhancements, blurring effects and weak edges which are the primary outcomes of processing by sharp enhancement or denoising filters. This paper proposes a full-reference perceptual quality estimation measure for color images incorporating the known characteristics of the HVS. The proposed evaluation model provides an objective assessment of the image by taking into account the error sensitivity, structural distortion and edge distortion. It uses the RGB model for color images and empirically combines the above effects on each of the color plane. The remaining part of this paper is structured as follows: Section II describes the method for estimation of degree of distortions and the proposed perceptual quality estimation measure. Details of the experimental procedure and the obtained results are discussed under section III. Finally the conclusions are drawn at the end.

II. Proposed Evaluation Model

Image degradations can be modeled as losses in the perceived structural content in addition to the error sensitivity estimation (as in Minkowski error metric [16]). There are still numerous structural changes which cannot be captured by the above estimators; one of them is edge distortion, a primary outcome of the images restored by sharp non-linear filters removing high density noise contaminations at the cost of weak edges and losses of finer details from the image. The motivation for our work is to develop a more precise evaluation framework for degraded images by modeling distortion in terms of error sensitivity along with structural and edge distortions for quality measurement. The proposed full-reference image quality measure (*Q*), performs evaluation as the weighted sum of the above mentioned distortions (i.e. error-sensitivity, structural and edge distortions) calculated for each of the R, G, B color planes separately. Figure 1 shows the block diagram for the proposed methodology for estimation of distortion in color images.

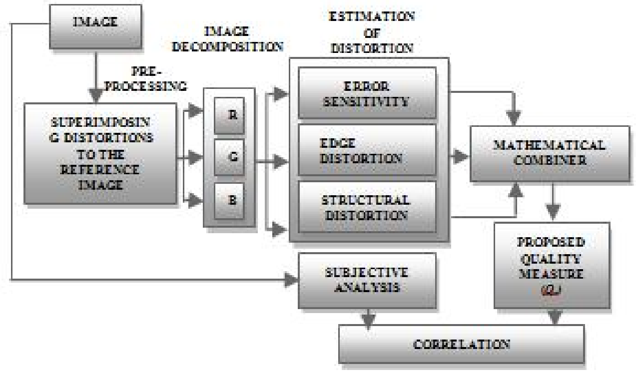
A. Measurement of Error

The distortion introduced in the image changes the pixel gray value which is mainly noticed by the HVS [12]. Thus, the error introduced in the three color components is calculated separately for quality measurement. If *x(i,j)* denotes the reference image and *y(i,j)*, the distorted image, then the error introduced in the R component is calculated as:

$$E_r = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[x(i,j)-y(i,j)\right]^2 \tag{1}$$

where: i & j are the pixel positions of the M x N image. Similarly, E_g and E_b are the error in G and B components, calculated as in (1). Overall error is estimated by average of the error for R, G and B components and hence, $PSNR_E$ *(dB)* due to error is given as:

$$PSNR_E = 10\,log_{10}\left[3/\left(E_r+E_g+E_b\right)\right] \tag{2}$$



B.Measurement of Structural Distortion

Human eyes mainly extract the structural changes from the viewing field, so structural distortion is an important parameter for quality assessment. Measurement of this distortion in the proposed work is performed by dividing the image in equal size PROPOSED QUALITY METRIC and non overlapping square regions, along with the calculation of the mean, maximum and minimum pixel values in each region. As it is known that the maximum and the minimum pixel values decide the contrast level of the region, thus the proposed measure also estimates the contrast change in the distorted image and hence can be used to evaluate the quality in coherence with HVS. Structural distortion S_r, S_g and S_b for the three color components R, G and B respectively can be calculated as:

$$S_r = \frac{1}{N}\sum_{i=1}^{N}\left\{0.5\left[Xa_i - Ya_i\right]^2 + 0.25\left[Xp_i - Yp_i\right]^2 + 0.25\left[Xb_i - Yb_i\right]^2\right\} \tag{3}$$

where: Xa_i, Xp_i, Xb_i and Ya_i, Yp_i, Yb_i denote the mean, maximum and minimum pixel values for the reference and the distorted image respectively. N is the number of regions in which the image is divided. The overall structural distortion is the mean of the structural distortion of the three color components and hence $PSNR_S$ *(dB)* of the structural distortion is given as:

$$PSNR_S = 10\,log_{10}\left[3/\left(S_r+S_g+S_b\right)\right] \tag{4}$$

C. Measurement of Edge Distortion

Edges denote the sites where there is an abrupt change in the pixel value or texture. A distorted image with very similar edges to the reference image gives a very high perceptual quality for HVS, although PSNR and SSIM produce an opposite result [9]. Canny edge detection algorithm [17] is used in this work as it is well-known for its large values of signal-to-noise ratio and high precision. The edge distortion for R component is given as:

$$ED_r = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[x_e(i,j)-y_e(i,j)\right]^2 \tag{5}$$

where: $x_e(i,j)$ and $y_e(i,j)$ denotes the original and distorted edge maps of the image. Similarly, ED_g and ED_b can be calculated for G and B component. Hence the overall $PSNR_{ED}$ *(dB)* of the edge distortion can be given as:

$$PSNR_{ED} = 10\,log_{10}\left[3/\left(ED_r+ED_g+ED_b\right)\right] \tag{6}$$

DOI: 01.IJSIP.03.01.11

D. Proposed Quality Estimation Measure

The proposed quality measure (*Q)* in *dB* is the weighted sum of different PSNR (in *dB*) components based on error, structural and edge distortion obtained from (2), (4) and (6) respectively.

$$Q = \alpha PSNR_E + \beta PSNR_{ED} + \gamma PSNR_S \tag{7}$$

where: α, β, and γ coefficients have different values, empirically determined for different types of distortions. Introduction of different coefficients in (7) is being done as each type of distortion has a different effect on error, structural and edge distortion respectively. Thus, these variable weights make the proposed quality measure a more sensitive and versatile towards specific type of distortions. In order to make the proposed measure distortion specific, distortions are divided into 3 categories. Distortions specific to noise attacks are grouped under category-I. Contrast distortions and degradations due to rotation are considered under the category-II where as category-III groups the blurring and compression artifacts. The quality measure *(Q)* in *dB* can be defined as per the given matrix:

$$\begin{bmatrix} Q_I \\ Q_{II} \\ Q_{III} \end{bmatrix} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.35 & 0.4 & 0.25 \\ 0.35 & 0.25 & 0.4 \end{bmatrix} \begin{bmatrix} PSNR_E \\ PSNR_{ED} \\ PSNR_S \end{bmatrix} \tag{8}$$

where: Q_I, Q_{II} *and* Q_{III} are the quality measure in *dB* for the distortion category-I, II and III respectively. In case of category I, noises contaminations deform the edges more severely than the others factors in the image. So, a higher coefficient is assigned to the PSNR component of edge distortion. The others two factors have an average affect on the image. Category-II considers contrast change and rotation in the image that only modifies the pixel values without causing any noticeable edge and structural distortion. Hence, a comparatively higher coefficient is assigned to the PSNR component for error sensitivity unlike category I. Category-III includes blurring and compression, which causes large structural modification in the image. This justifies for the choice of the higher coefficients for the PSNR component of error and structural distortion respectively. A generalized version of the proposed quality measure known as (Q_0) in *dB* is given in (9), which is independent of the distortion type.

$$Q_0 = 0.32 PSNR_E + 0.38 PSNR_{ED} + 0.3 PSNR_S \tag{9}$$

III. Experimental Results and Discussion

A. Experimental Procedure

The input reference image used for the experiments in this work is standard Lena image of size 256 x 256. As shown in figure 1, the early transformations (like normalization, gray conversion etc.) are applied to the input image in the pre-processing block. Different types of distortions are then superimposed on the pre-processed image; these include noise contaminations (salt & pepper noise, speckle noise and Gaussian noise), blurring (motion & Gaussian blur), contrast change, compression and rotation.

Figure 2 shows the different simulated distortions applied to the pre-processed input reference image. The color image is then decomposed into its three components (R, G & B color planes) for further processing. The PSNR component for error sensitivity, structural and edge distortions are then calculated for all the components of the image as described in (1)-(6). Finally, the obtained results are empirically combined in the mathematical combiner to yield the quality measure *(Q)* in *dB as* in (7)-(8), which specifically quantifies the effect of different types of simulated distortion. Table I gives the comparison of the obtained values of proposed measure (*Q)* in *dB* along with the conventional PSNR (*dB*) for each category of distortions.

Table I. Evaluation results of conventional PSNR and proposed quality metric (Q).

Category/Figure		PSNR*(dB)*	*Q(dB)*	Subjective Score
I	(a)	18.6320	10.2196	0.4298
	(b)	14.5678	8.1119	0.3051
	(c)	11.5050	6.9081	0.1653
	(d)	22.8177	13.6865	0.5378
	(e)	19.9960	11.7725	0.4400
	(f)	17.2710	10.0147	0.3589
	(g)	19.3270	11.1124	0.4109
	(h)	15.7917	8.9415	0.3186
	(i)	15.0232	8.5465	0.2766
	(j)	13.1031	7.5654	0.2126
II	(k)	21.8665	17.6067	0.8207
	(l)	18.4490	13.7423	0.7815
	(m)	15.5080	11.3577	0.7004
	(n)	11.7482	10.4492	0.6766
	(o)	9.6321	9.0090	0.6872
III	(p)	26.1186	22.7422	0.5679
	(q)	22.7101	19.7818	0.3483
	(r)	21.6101	18.9374	0.2885
	(s)	22.2402	19.4845	0.3509
	(t)	21.2416	18.6720	0.3058
	(u)	20.0157	17.7129	0.2285
	(v)	19.1206	17.0083	0.2091
	(w)	21.3777	18.8795	0.3293
	(x)	20.3777	18.0888	0.2376

DOI: 01.IJSIP.03.01.11

Figure 2. Reference image simulated with different types of distortion. (a) - (c) Salt & Pepper noise. (d) - (f) Speckle noise. (g) - (j) Gaussian noise. (k) -(m) Contrast stretching. (n), (o) Rotation. (p) - (r) Gaussian blurring. (s) - (v) Motion blurring. (w), (x) Compression.

B. Subjective Evaluation

In this paper the efficiency of the objective quality measure is adjudged by the subjective evaluation model. In the process of subjective evaluation 42 images have been assessed upon their visual quality by a group of 150 observers. These set of images have been distorted by varying levels of specified distortions. The quality scores given by the observers were recorded on a quality scale of 0 to 1 (where 0 means worst and 1 means best). These subjective scores were directly related to the quality of the image i.e. higher the score better is the visual quality of the image.

C. Processing of Subjective Scores

Let s_{ij} denotes the subjective score assigned by the observer i to the image j. In this paper the subjective evaluation is based on the *DMOS* i.e. difference mean opinion score, which basically takes in account the difference between the scores of reference and the test image. So, the difference score d_{ij} has been calculated by subtracting the subjective score assigned by the observer to the test image from the scores assigned to the reference image.

$$d_{ij} = s_{ij_{ref}} - s_{ij} \tag{10}$$

The difference scores that are zero in the case of reference image have been removed for further processing. These difference scores are converted to the Z-scores [18] by using the following equations:

$$\mu_i = \frac{1}{N} \sum_{j=1}^{N} d_{ij} \tag{11}$$

$$\sigma_i = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} (d_{ij} - \mu_i)^2} \tag{12}$$

$$z_{ij} = \frac{d_{ij} - \mu_i}{\sigma_i} \tag{13}$$

Where N is the number of test images seen by the observer. These Z-scores are then linearly scaled between [0, 1] in order to develop a proper evaluation model. In our experiment all the Z-scores lie within the range of [-4, 5]. Hence rescaling was accomplished by linearly mapping the range [-4, 5] to [0, 1] using:

$$z'_{ij} = (z_{ij} + 4) / 9 \tag{14}$$

Finally the Difference Mean Opinion Score (*DMOS*) has been calculated as the average of all the Z-scores for the respective image as:

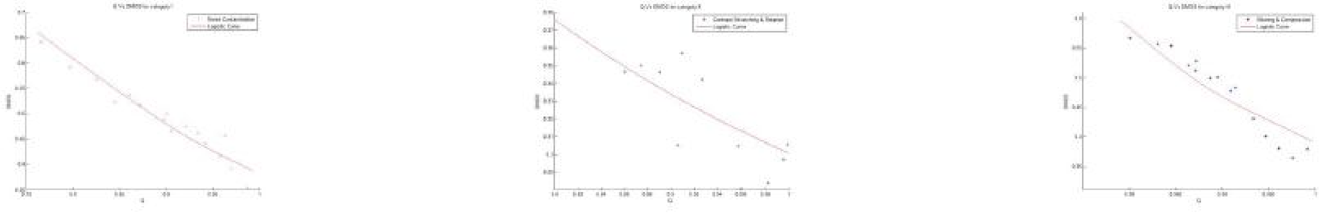
$$DMOS = \frac{1}{M} \sum_{i=1}^{M} z'_{ij} \tag{15}$$

where: M is the total no of observers in the subjective evaluation [19].

D. Comparison of Results

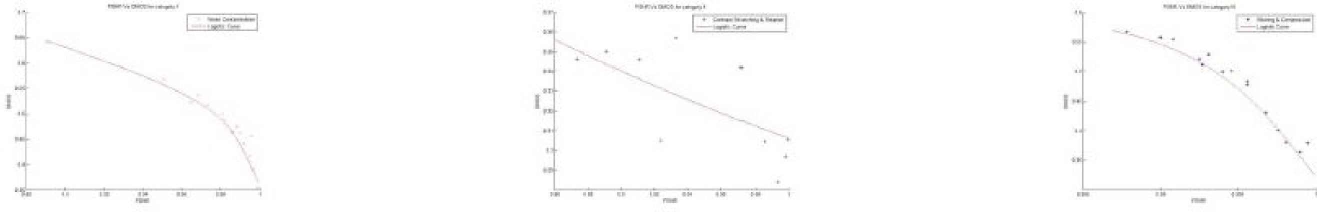
In order to evaluate the performance of the proposed quality measure *(Q),* Pearson correlation coefficient is used as a tool to estimate the correlation of Q with respect to different HVS based metrics like Edge Performance Index *(EPI),* Structural Correlation *(SC),* Mean Absolute Error *(MAE)* [20], CONTRAST (c(x,y)) [21], and *DMOS.* The results in table I show that there is a large difference in values of conventional PSNR as compared to Q, for the images with approximately same visual quality. It can be inferred from the results stated in table II that Q is more efficient in measuring the degradation for all the distortion categories as per the HVS characteristics. From the results stated in table III, it can be seen that Q in addition of being a better HVS based quality measure, also have the capability to assess specific types of distortions like edge, structure, contrast changes and error in the image as the results have very high correlation coefficients. All the stated results validate the effectiveness of proposed quality measure in measuring the degradation in the image for the various distortion types. The effectiveness of the quality measure Q can also be seen from the scatter plots shown in fig. 3, 4 and 5. In fig. 3(a), variation of the values of Q with the *DMOS* for noise contamination (category I) depicts that for higher values of Q the *DMOS* score is low i.e. the *DMOS* values decreases with the increase in the visual quality of the image because of lower noise contamination. The visual quality of the image degrades with the increase in the degree

DOI: 01.IJSIP.03.01.11



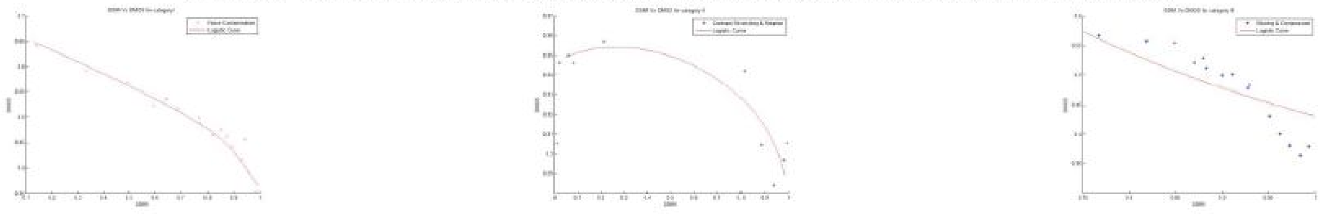
(a) (b) (c)

Figure 1. Scatter plot of Q Vs *DMOS* for (a) Category I. (b) Category II. (c) Category III.



(a) (b) (c)

Figure 2. Scatter plot of PSNR Vs *DMOS* for (a) Category I. (b) Category II. (c) Category III.



(a) (b) (c)

Figure 3. Scatter plot of SSIM Vs *DMOS* for (a) Category I. (b) Category II. (c) Category III.

of distortions like compression, blurring, contrast change etc, which is also justified from the variation of *Q* shown in fig. 3 (b) and (c). For the same type and degree of distortion, the values of PSNR and SSIM have also been plotted in fig. 4 and 5 but their values does not follow any regular patter of variation with the subjective *DMOS* score showing their inefficiency in evaluating the results as per the HVS characteristics.

TABLE II. PEARSON CORRELATION BETWEEN DIFFERENT QUALITY MEASURES AND SUBJECTIVE EVALUATION

Category	Distortion Types	PSNR	SSIM	Q_P
I	Noises	0.9069	0.9470	0.9712
II	Contrast Stretching & Rotation	0.7380	0.6785	0.8052
III	Compressing & Blurring	0.9543	0.9325	0.9705

CONCLUSION

This paper presents a new perceptual quality estimation model for color images based on HVS. The proposed model accounts for all the types of degradations that severely affect the visual quality of the image. It evaluates the quality measure *Q* as weighted sum of PSNR components due to different distortions for each color planes. Different sets of weights have been used according to the category of the distortion and their values are based on the effect of the distortion on the visual quality of the image i.e. how it affects some of the common properties of the HVS like error, structural and edge information of the image. Weighted coefficients used in the quality measure also strengthen its quality measurement ability. The correlation coefficients calculated for the subjective evaluation results prove the efficiency of proposed quality measure and its consistency with the HVS characteristics. The simulation results also show strong correlation with several other well known distortion measures like *EPI, SC, MAE* and *CONTRAST*.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," *2nd eds., Prentice Hall*, Upper Saddle River, New Jersey, 2001.

[2] N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, M. Carli, and F. Battisti, "Color Image Database for Evaluation of Image Quality Metrics," *IEEE Signal Process. Society 2008 Int. Workshop on Multimedia Signal Process.*, Australia, pp. 403-408, Sept 2008.

[3] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.*, vol. 4, (Orlando), pp. 3313–3316, May 2002.

[4] Z. Wang and A. C. Bovik, "Mean Squared Error: Love It or Leave It?," *IEEE Signal Process. Magazine*, pp. 98-117, Jan. 2009.

[5] Z. Wang, L. Lu, and A.C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process. Image Commun.*, vol. 19, no. 2, pp. 121–132, Feb. 2004.

[6] Z. Wang, A. C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli, "Image Quality Assessment: From Error Measurement to Structural Similarity," *IEEE Trans. Image Process.*, Vol. 13, no. 4, pp. 600-612, April 2004.

[7] E. Wharton, K. Panetta and S. Agatan, "Human Visual System Based Similarity metrics," in *Proc. of IEEE Conf. on System, Man and Cybernetics*, pp. 685-690, Oct. 2008.

DOI: 01.IJSIP.03.01.11

[8] W. Yang, L. Wu, Ye fan and Z. Wang, "A Method of Image Quality Assessment Based on Region of Interest," in *Proc. of the 7th World Congress on Intelligent Control and Autom.*, Chongqing, China, pp. 6840-6843, June 2008.

[9] Wei Fu, Xiaodong Gu*, Yuanyuan Wang , "Image Quality Assessment Using Edge and Contrast Similarity," in *Proc. of Int. joint Conf. on Neural Networks*, pp. 852 -855, 2008.

[10] B. Wang, Z. Wang, Y. Liao and X. Lin, " HVS-based structural similarity for image quality assessment," in *Proc. of 9th Int. Conf. on Signal Process.*, Beijing, pp. 1194–1197, December 2008.

[11] Ho-Sung Han,Dong-O Kim and Rae-Hang Park, " Gradient Information- Based Image Quality Metric," *Consumer Electronics (ICCE), 2010 Digest of Technical Papers*, pp. 361-362, Feburary 2010.

[12] Chen Yutuo, Wu Meijie, Fei Yongchao, "Evaluation Method of Color Image Coding Quality integrating Visual Characteristics of Human Eye," in *Proc. of 2nd Int. Conf. on Education Technology and Computer*, Shanghai, China, Vol. 2, pp. 562 -566, June 2010.

[13] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *http://www.vqeg.org/*, Mar. 2000.

[14] J.-B. Martens and L. Meesters, "Image dissimilarity," *Signal Processing*, vol. 70, pp. 155–176, Nov. 1998.

[15] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Commun.*, vol. 43, pp. 2959–2965, Dec. 1995.

[16] T. N. Pappas and R. J. Safranek, "Perceptual criteria for image quality evaluation," *in Handbook of Image and Video Proc. (A. Bovik, ed.)*, Academic Press, 2000.

[17] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[18] A. M. van Dijk, J. B. Martens, and A. B. Watson, "Quality assessment of coded images using numerical category scaling," in *Proc. SPIE -Advanced Image and Video Commun. and Storage Technologies*, 1995.

[19] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "Study of Subjective and Objective Quality Assessment of Video," *IEEE Trans. Image Process.*, Vol. 19, no. 6, pp. 1427-1441, June 2010.

[20] Neha Tripathi, Mohd. Faizan Wahidi, Ashwani Gupta, and Vikrant Bhateja, "A Novel Spatial Domain Image Quality Metric," in *Proc. of (IEEE) 2011 World Congress on Engg. & Technology (CET-2011)* Shanghai: Signal & Information Process., pp. 902-905 , Oct. 2011.

[21] W.M. Morrow, R.B. Paranjape, R.M. Rangayyan and J.E.L. Desautels, "Region Based Contrast Enhancement of Mammograms," *IEEE Trans. on Medical Imaging, vol. 11, no. 3,* pp.392-406, 1992.

Table III. Pearson Correlation between different quality measures and EPI , SC, MAE and CONTRAST(c(x,y))

	EPI			SC			MAE			CONTRAST		
Category/Distortion Types	PSNR	SSIM	Q_P	PSNR	SSIM	Q_P	PSNR	SSIM	Q_P	PSNR	SSIM	Q_P
I -Noises	0.7862	0.8643	0.9441	0.9769	0.9849	0.9548	0.8831	0.8771	0.8382	0.9928	0.9610	0.8861
II-Contrast Stretching & rotation	0.9271	0.9701	0.9551	0.9839	0.9848	0.9324	0.9425	0.9878	0.9813	0.0918	-0.0842	0.2871
III-Compressing & blurring	0.7776	0.7401	0.8337	0.9998	0.9953	0.9953	0.9901	0.9759	0.9958	0.8658	0.9090	0.8470

DOI: 01.IJSIP.03.01.11